Counting Rational Points over Finite Fields

UMN/IMA.
9/18/06.

$\mathbb{F}_q$: finite field of q elements. $q = p^k$.

X : affine hypersurface $f(x_1, \cdots, x_n) = 0 \hookrightarrow \mathbb{A}^n$

$N(f) := \#\{(x_1, \cdots, x_n) \in \mathbb{F}_q^n \mid f(x_1, \cdots, x_n) = 0\}$

Problem: Compute $N(f)$ efficiently.

Clearly, $0 \le N(f) \le q^n$.

I. Sparse input: $f(x_1, \cdots, x_n) = \sum_{j=1}^{m} a_j x_1^{v_{j1}} \cdots x_n^{v_{jn}}$, $0 \le v_{jk} \le q-1$.

$a_j \in \mathbb{F}_q^*$.

Sparse input size. $O(mn \log q) \ge$ output size $n \log q$.

Sparse poly time: $O(mn \log q)^c$, c = constant.

II Dense input (of poly of deg $\le d$).

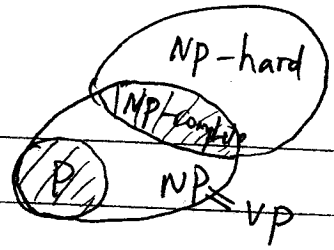
$$f(x_1, \cdots, x_n) = \sum_{0 \le i_1 + \cdots + i_n \le d} a_{i_1 \cdots i_n} x_1^{i_1} \cdots x_n^{i_n}$$

$a_{i_1 \cdots i_n} \in \mathbb{F}_q$.

dense input size. $O\left(\binom{d+n}{n} \log q\right) \ge$ output size $n \log q$

dense poly time: $O\left(\binom{d+n}{n} \log q\right)^c$.

I. Sparse Case.



1. Complexity.

$$P \leq NP\text{-complete} \leq NP\text{-hard}$$

Prop. Computing $N(f)$ is NP-hard.

Pf. Take $f = a_1 x_1^{q-1} + \cdots + a_n x_n^{q-1} - a_0$.

Then $N(f) > 0 \iff \exists$ non-empty subset

$$\{a_{i_1}, \cdots, a_{i_k}\} \subseteq \{a_1, \cdots, a_n\}$$

s.t. $a_{i_1} + \cdots + a_{i_k} = a_0$.

But this $\mathbb{F}_q$-subset sum problem

is $\in \begin{cases} P, & \text{if } p=2 \\ NP\text{-complete}, & \text{if } p>2. \end{cases}$

Cor. For each fixed $p>2$, $\Rightarrow$ computing $N(f)$ is NP-hard.

2. Complexity of modular counting.

Problem: Given integer $r > 1$. Compute $N(f)$ mod r.

If easy for many small r, $\Rightarrow$ By Chinese Remainder Thm,

easy for $N(f)$.

(WLOG, can assume r is a prime power).

Prop (Goplan-Guruswami-Lipton, 06). For each fixed $r > 1$, computing $N(f) \bmod r$ is NP-hard.

⇑

Prop. Let $q = p^h$

1) If $r \neq p^b$, $\Rightarrow$ Computing $N(f) \bmod r$ is NP-hard

2) If $r = p^b$, $\Rightarrow$ Computing $N(f) \bmod r$ is NP-hard if either $p \geq 2n$ or $h \geq 2n$ ($\Leftarrow$ q is "large").

Cor. Exponential dependence on $\{p, b, h\}$ cannot be avoided in computing $N(f) \bmod r$ (within current knowledge).

3. Algorithms for modular counting.

Thm (GGL, 06). Let $q = p^h$. Then $N(f) \bmod p^b$ can be computed in time $O(n m^{2^{8b}}) = O(n m^{2p^h \cdot b})$.

($\Rightarrow$ poly time if q and b are fixed) but doubly exponential in h.

Thm (W., 06). Let $q = p^h$. Then $N(f) \bmod p^b$ can be computed in time $O(n (qm)^{(h+b)p})$

($\Rightarrow$ singly exponential in each of $\{p, b, h\}$).

Question. Can one compute $N(f) \bmod p^b$ in time $O(n m^{\varepsilon(h+b)p})$ for some $0 < \varepsilon < 1$,

Two Pf of Thm:

1) Gauss sum, Stickelberger thm, Gross-Koblitz formula via p-adic Γ-function.

2) Dwork's p-adic method.

4. An example (from coding theory)

Let $1 \le m \le n$. Let V be the symmetric variety

$$\begin{cases} \sum_{i=1}^{n} x_i = b_1 \\ \sum_{1 \le i_1 < i_2 \le n} x_{i_1} x_{i_2} = b_2 \\ \vdots \\ \sum_{1 \le i_1 < i_2 < \cdots < i_m \le n} x_{i_1} \cdots x_{i_m} = b_m \end{cases} \qquad b_i \in \mathbb{F}_q \text{ given}$$

$\Leftrightarrow$ Newton's formula (if $p > m$)

$$\begin{cases} x_1 + \cdots + x_n = c_1 \\ x_1^2 + \cdots + x_n^2 = c_2 \\ \vdots \\ x_1^m + \cdots + x_n^m = c_m. \end{cases}$$

Let $N(V) = \#$ of $\mathbb{F}_q$-rational pts on V.

Problem: Is it NP-complete to determine if $N(V) > 0$?

(using the input size $O(mn \log q)$).

Conj. This problem $\in P$ for small m.

If $m=1$, $\Rightarrow$ $N(V) = q^{n-1} > 0$.

For small m, using RH over $\mathbb{F}_q$, $\Rightarrow$

Thm Let $m < \frac{n}{2}$. If $q > (m-1)^2 + \frac{4m}{n-2m}$, $\Rightarrow N(V) > 0$.

($m=2$, $\Rightarrow$ ok. ✓)

($m=3$, $\Rightarrow$ can assume $q = 2, 3, 4$)

For general m and q, much more than the Weil conjectures is needed!

II Dense case.

1 Zeta function

Def. The zeta function of $X/\mathbb{F}_q$ is

$$Z(X, T) = \exp\left(\sum_{k=1}^{\infty} \frac{T^k}{k} \#X(\mathbb{F}_{q^k}) \right)$$

$$= \prod_{\substack{x \in |X| \\ \text{closed pts}}} \frac{1}{1 - T^{\deg(x)}} \in 1 + T\mathbb{Z}[[T]].$$

Thm 1) (Dwork) $Z(X, T) \in \mathbb{Q}(T)$

2) (Bombieri) Total $\deg(Z(X,T)) \leq (4d+9)^{n+1}$.

Output size $(Z(X,T))$ $\geq$ dense input size $O(\binom{d+n}{n} \log q)$

$\sim$ if n is fixed.

Problem. Compute $Z(X, T)$ efficiently ($\Leftarrow$ Compute $\#X(\mathbb{F}_{q^k})$ $\forall k = 1, 2, \cdots$)

2. Cohom. formulas

$$Z(X,T) = \prod_{i=0}^{2\dim(X)} \det\left(I - \mathrm{Frob}_q T \,\middle|\, H_c^i(X)\right)^{(-1)^{i-1}}$$

ℓ-adic coh $H_c^i(X)$, $\ell \neq p$. Grothendieck, not effective in general

p-adic coh $H_c^i(X)$, $\ell = p$. Dwork, effective!

and efficient if p is small.

3. p-adic algorithms.

Thm (W, 97). $Z(X,T)$ mod p can be computed in time

$$O\left(p \binom{d}{n} \log q\right)^c, \Rightarrow \text{poly time in dense input}$$

if p is small.

99. same for $Z(X,T)$ mod p^b if p^b is small.

Thm (Lauder–W, 02). $Z(X,T)$ can be computed in time

$$O(p\, d^n \log q)^{O(n)} \Rightarrow$$

($\Rightarrow$ poly time if p small and n fixed)

in dense input

$d^n \sim \binom{n+d}{n}$ if n fixed.

Note. $Z(X,T)$ cannot be computed in poly time in dense input size for large n as output size $d^n \log q > \binom{d+n}{n} \log q$.

Thm (Lauder, 04) If X is a smooth proj hypersurface of degree d not divisible by p, $\Rightarrow$

$Z(X, T)$ can be computed in time $O(p d^n \log q)^{O(1)}$

($\Rightarrow$ poly time in ~~(dense input + output)~~ if p is small.

Question: Is this thm true for arbitrary singular ~~hypersurface~~ of deg d ?